P1 1199966

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

July 27, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: ***60/464,617***
FILING DATE: ***April 23, 2003***
RELATED PCT APPLICATION NUMBER: *PCT/US04/12363*

REC'D 30 JUL 2004
WIPO PCT

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

T Lawrence

T. LAWRENCE
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

60464617 04/22/03

J1031 U.S. PTO
04/23/03

J1051 U.S. PTO
60/464617

PTO/SB/16 (10-01)
Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No.

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Richard H.	Lyon	Belmont, MA
David L.	Bowen	Cambridge, MA
Gladys L.	Unger	Belmont, MA

☐ *Additional inventors are being named on the ______ separately numbered sheets attached hereto*

TITLE OF THE INVENTION (500 characters max)

The Accelophone - A transducer for improved speech privacy and immunity from ambient noise

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ Customer Number ______ *Type Customer Number here* → *Place Customer Number Bar Code Label here*

OR

[X] Firm or Individual Name	RH Lyon Corp
Address	691 Concord Avenue
Address	
City	Cambridge — State: MA — ZIP: 02138
Country	USA — Telephone: 617-864-7260 — Fax: 617-864-0779

ENCLOSED APPLICATION PARTS (*check all that apply*)

[X] Specification *Number of Pages* 6
[X] Drawing(s) *Number of Sheets* 4
☐ Application Data Sheet. See 37 CFR 1.76
☐ CD(s), Number
☐ Other (specify)

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

[X] Applicant claims small entity status. See 37 CFR 1.27.
[X] A check or money order is enclosed to cover the filing fees
☐ The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number:
☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE AMOUNT ($): 80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
[X] No.
☐ Yes, the name of the U.S. Government agency and the Government contract number are: ______

Respectfully submitted,

SIGNATURE Richard H Lyon

Date 04/22/03

TYPED or PRINTED NAME Richard H. Lyon

REGISTRATION NO. *(if appropriate)*
Docket Number:

TELEPHONE 617-864-7260

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

[redacted]

A Transducer for Telephone Privacy and Reduced Interference from Background Noise

This document describes a new transducer for the sensing of sounds from a talker using the acceleration of the air. Further optimization of this acceleration is accompanied by reduction of the portion of the sound energy that escapes from the regions around the transducer. The result is a high sensitivity transducer with increased privacy as a result of the reduction in radiated sound, with significant advantages for use in communication systems, especially cell phones and in an office environment.

Three design problems are inherent in telephonic and other communications systems. These difficulties include: (1) sensitivity to acoustical background noise that interferes with understanding, (2) radiation of sound to others allowing them to overhear private conversation, and (3) sensitivity to wind noise produced primarily by locally generated turbulence. Problems (1) and (2) are closely related although not identical. In the case where the user's lips are very close to the microphone these two concerns may be related through reciprocity – if sound is well received from a given direction it will, by reciprocity, be well radiated back into that same direction. Military and industrial systems in general have the first problem because they often operate in regions of high noise level. Cell phones and other telephonic communication systems, for which privacy is an issue, often have the latter problem.

Noise due to turbulence is usually addressed by surrounding the transducer with a windscreen. Windscreens are commonly made from a porous (open cell) plastic foam material. These windscreens can be effective, but their size can be an issue, and in a high wind they lose their effectiveness. Microphone arrays can also reduce sensitivity to local pressure fluctuations produced by turbulence, but at a penalty related to transducer size, complexity, and cost.

This invention is concerned with the realization of a transducer that measures and maximizes the *acceleration* of air particles in front of the mouth, in contrast with conventional microphones which measure sound pressure. We call this device the AcceloPhone™. It is configured to be sensitive to sound produced only by the user, and by its design minimize sound waves that would otherwise radiate away from the transducer and heard by others. The layout of the transducer's components is illustrated in Figure 1. The technology is explained later in this document. Once constructed, the actual integration of the AcceloPhone™ into telephone and other communications systems can be accomplished easily.

The AcceloPhone™ consists of two or more closely spaced microphones and a loudspeaker. The idea of the AcceloPhone™ is to employ a loudspeaker to draw in volume velocity that is produced at the lips and nose of a talker. When the microphones are close together (within about one-sixth of a wavelength of sound at the highest frequency of interest), then inertial effects of the air (represented by an acoustic mass) dominate the pressure difference between the microphones. If the loudspeaker is drawing in volume velocity at the same rate as the talker is producing volume velocity, then the pressure gradient is maximized and the pressure, and consequently, the compression of the air is minimized. Therefore, the sound produced, i.e. the sound pressure, will be very weak since volume velocity does not "escape" the transducer to produce sound away from the talker. A demonstration of the AcceloPhone™ has been constructed and a photograph of this unit is shown in Figure 2.

Although the pressure and air compression are minimized, the air in the immediate region between the talker and the loudspeaker is accelerated and that acceleration is proportional to the pressure gradient. It is the pressure gradient, or equivalently the air acceleration, that the AcceloPhone™ measures and passes on as the signal to be communicated. The measurement of pressure gradient itself is not new and is routinely incorporated into the "acoustical intensity probes" in everyday use by acoustical engineers. It is also the basis of operation of the "ribbon microphones" in common use in radio broadcasting for many years. We note that this pressure gradient signal is measurable even if the air were incompressible, and that is the key to its reduced sensitivity to ambient sounds and its minimization of radiated sound. The

[redacted]

Accelophone™ deliberately minimizes the sound produced by the talker while it enhances the oscillatory flow of air over the microphone pair. Although pressure gradient microphones also measure the acceleration of the air, they do not maximize acceleration and minimize compression by the incorporation of a local loudspeaker as the Accelophone™ does.

One characteristic of a transducer that is based on measuring pressure gradient rather than pressure is the frequency coloration that may be introduced. Since the gradient introduces a spatial derivative, high frequencies are emphasized. However, the air velocity near the mouth is a time integral of the pressure, and since the acceleration of the air is the time derivative of the velocity, the acceleration has the same frequency characteristic as the pressure. A second question may arise; if the loudspeaker is absorbing the volume flow generated by the talker, will the talker be able to hear his/her own voice? This is not a problem, since much of what we hear of our own speech is due to tissue and bone conduction, not to the sound traveling to our ears through the air.

To increase noise immunity from the turbulent airflow, a shroud, such as the one shown in Figure 1, can be incorporated into the realization of the AcceloPhone™ device. The shroud can be optimized to reduce the effects of turbulence. A porous foam windscreen can also be incorporated into this transducer.

Analysis and Operation of the Accelophone

A sketch of one realization of the AcceloPhone™ system is shown in Figure 1. The transducer elements consist of a pair of microphones and a small loudspeaker. The speaker is driven by a signal proportional to the *difference* in the microphone outputs in such a way as to *minimize the pressure as measured by the sum* of the microphone outputs. The acceleration is given by the formula $a(t) \approx (p_1 - p_2)/\rho\Delta x$, where ρ is the density of the air and Δx is the distance between the microphones (this relationship is altered when turbulence is present - see the discussion below). If an array of more than two microphones is employed, their outputs are to be combined in such a way so that the total pressure is to be minimized and the pressure gradient is simultaneously maximized.

The system therefore maximizes the acceleration of the air in the region between the lips and the loudspeaker above that which would be sensed by an ordinary velocity or pressure gradient microphone. At the same time, it minimizes the compression of the air in that region.

The concept can be demonstrated by three areas of activity: (1) modeling the acoustical processes involved, (2) incorporating the appropriate signal processing algorithms, and (3) building a prototype and testing for immunity to ambient acoustical noise and minimizing the sound radiated away from the AcceloPhone™.

Analytical model of the acoustical processes

The simplest acoustical model of the talker using the AcceloPhone™ treat the system as a dipole since the loudspeaker will draw in volume velocity equal to that produced by the lips and nose of the talker. This increases the acceleration of the airflow and reduces the pressure produced by the talker. For purposes of discussion we use the two microphone arrangement of Figure 1, but similar and potentially improved results are achievable with an array of microphones. Within a 6^{th} of a wavelength (about 1 inch at 2000 Hz, the upper range of frequency for speech) incompressible terms in the flow field dominate, and the radiated pressure has the dipole directionality of $|\cos\theta|$, which reduces the radiation to the surrounding area for privacy. Since the acceleration in the region of the microphones is proportional to the pressure gradient, acceleration has a directivity of $|\cos^2\theta|$, which reduces the sensitivity to background noise further. Relative directivities for a dipole sensor based on pressure and one based on acceleration for the two microphone arrangement are shown in Figure 3.

[redacted]

These simple estimates are readily refined with more detailed calculations for the actual geometries involved using computational methods to get more detail on the expected benefits, both in terms of privacy and insensitivity to ambient sounds. Independent of our software choice we employ a realistic model for the acoustics of the AcceloPhone™ system.

The acoustical inputs to the AcceloPhone™ are the volume velocity from the talker's lips and nose, U_L, and the volume velocity from the loudspeaker, U_S, as indicated in Figure 1. The volume velocity, U_S is determined by the voltage applied to the loudspeaker, V_S. For the purpose of this discussion, we estimated the pressure gradient using only two microphones, as shown in Figure 1. The pressures at locations (1) and (2) are sensed by microphones at those locations which then produce signals proportional to p_1 and p_2. The purpose of the model is to develop the functional relationships among these variables for guidance in design optimization. A model can provide a good indication for the directions that system parameters should be changed for improved behavior.

The greatest need for the model is in dealing with the geometry of the space between the talker and the loudspeaker. If a shroud is present, as indicated in Figure 4, then the acoustics are different than if there is no shroud. The acoustical model has to deal with that option. If the spacing between the two microphones is quite small, then compression in the air between them can be neglected and the element H_{12} in Figure 1 becomes a simple acoustical mass, the value of which will still depend on the shroud geometry. The spaces between the talker and microphone #1 and between microphone #2 and the loudspeaker are more complicated and will need the assistance of a computational model for definition in the design.

Experience with telephonic transmission indicates that this system needs to be effective over a frequency range from about 200 to 2000 Hz. The loudspeaker needs to maximize the acceleration at the microphone array until the pressure is minimized and must therefore be able to react to signals within a fraction of the period of the highest frequency of interest. Delays in the system, electrical, mechanical, and acoustical, need to be minimized. The analytical model is to be used extensively for this minimization. For example, sound travels about 35 mm in 100 µsec. If the largest propagation delay we can tolerate is 200 µsec (0.4 periods at 2000 Hz), then the distance limit between the loudspeaker and the nearest microphone is limited to 70 mm. There is no such restriction on the distance between the talker and the microphones however. Delay between the time of actual speech production and its arrival at the AcceloPhone™ may affect privacy, but should not affect the immunity from ambient sound and sensitivity to speech from the talker.

The form of the final analytical model is shown in Figure 5. This diagram anticipates the structure of the model. Computational analysis can be used to quantify the elements shown. The boundary element acoustical model (BEMAP), and finite element algorithm (ALGOR) are example programs that can be used to represent the acoustics of this space. The principal use of the model will be to determine the effects of variations that are inherent in any constructed system on the performance of the system as a whole. For example, it is desirable to keep the spacing between the two microphones as small as possible. It is possible to minimize this distance if phase-matched microphones are used, but such microphones can be expensive, and other approaches may be necessary. The acoustical analysis must be carried out in conjunction with algorithm choices and experimental evaluations.

<u>Algorithm for maximizing acceleration and minimizing pressure – two microphone example</u>

The microphones will measure the pressure fields in their vicinity, p_1 and p_2. The acceleration is proportional to the pressure difference so the required voltage to the loudspeaker is given by $V_S=K(z)(P_1 - P_2)$, where K is chosen to minimize $|p_1+p_2|$. The exact form of K to achieve the minimization will depend on the loudspeaker and on the geometry of the AcceloPhone™. It may also depend on

[redacted]

the geometry of the user's face and other items that will vary from one situation to another. The acoustical model is used to account for this acoustical variability. Once K is determined, the loudspeaker voltage V_S is used to provide the signal desired: $\Delta P = K^{-1} V_S$. Alternately, the two microphone signals themselves may be used as the signal to be transmitted from the Accelophone™. This design step is necessary in practice to determine the best procedure for estimating K and the sensitivity of performance to variations of K.

The purpose of the microphones is primarily to measure the pressure gradient in the region between the talker and the Accelophone™ speaker. Since there must be a finite distance between the microphones, estimating the gradient can be improved by an increasing the number of microphones in a way that is well known from finite difference analysis. It is the estimation of the pressure gradient from microphone measurements that is part of the special character of the Accelophone™. A pair of microphones is adequate for an estimate, but a larger number may be used to improve the estimate.

The two-microphone arrangement is used here to demonstrate the principles. The acceleration of the air in the space between the two microphones, $a(t)$, is governed by the difference in the pressures, $\partial p / \partial x = -\rho a(t)$, or $a(t) \approx (p_1 - p_2) / \rho \Delta x$. The processing algorithm is used to maximize the pressure difference, $p_1 - p_2$, while minimizing the pressure sum, $p_1 + p_2$. To achieve this in the frequency domain, the voltage applied to the loudspeaker should be proportional to that difference, e.g., $V_S = K(P_1 - P_2)$, where the magnitude and phase functions of K are chosen to minimize the sum of complex amplitudes $P_1 + P_2$. The optimized acceleration, which is the output of the AcceloPhone™, is then readily calculated from the voltage V_S as noted above. We then have a choice for the output signal from the system; either derive it from the loudspeaker voltage or take it directly from the microphone array (the set of two microphones as shown in the figure, but an array of more than two microphones is possible).

When turbulence is present, the relationship between the pressure gradient and the acceleration is altered to become (for one dimensional inviscid flow), $\partial p / \partial x = -\rho a(t) - \partial(\rho u^2 / 2) / \partial x$, where $u = \int a \, dt$ is the velocity of the airflow. The new term in the equation is the convective acceleration and its presence means that the relation between pressure and acceleration assumed for the AcceloPhone™ is altered. In turbulence, the two terms on the right-hand side may be comparable in magnitude. However, since we measure pressure gradient it may be possible to back out a velocity estimate and "correct" for some of the turbulence effect. The consequences of this are not clear at this time, but it may mean that the AcceloPhone™ will always require some sort of windscreen for protection if airflow noise isa problem.

While the operation of the AcceloPhone™ is similar to an active noise canceller in minimizing the total pressure, unlike active noise canceller, the objective of AcceloPhone™ is to maximize the pressure gradient represented by Δp. If the sound does not come from the talker, but from some other direction, then reducing the total pressure will not maximize Δp, and the output from the system will be reduced, leading to the desired immunity from ambient sound.

Figure 6 indicates the conceptual operation of the optimization procedure. The combination of the talker and the loudspeaker, or the loudspeaker alone in a calibration mode, creates the pair of microphone outputs p_1 and p_2. The sum of these is p_{tot} and the difference is Δp. An initial estimate of K is made and a voltage $V_S = K \Delta p$ is applied to the loudspeaker and the summed pressure magnitude $|p_{tot}|$ is compared with a limit error value ε. A modification, ΔK, to the function K is then produced and a new voltage signal is sent out to the loudspeaker.

The analytical model is used to come up with a practical optimization approach. For example, Figure 7 illustrates a method for continuously adjusting the filter coefficients of K, based on a time-domain adap-

[redacted]

tive approach utilizing a variant of the normalized filtered-x LMS algorithm (other approaches, utilizing direct minimization / maximization of p_{tot} and Δp via modifications of K with appropriate constraints imposed, may also be possible if a detailed enough model is available). In general for such an approach, the pressure sum (in the form of $-p_{tot}/2$) would be used as the "error" signal input to the adaptive algorithm, while the pressure difference would be used as the "reference" (x) signal input. In order to compensate for delays and other transfer path distortions introduced by the physical systems between the output of the filter K and the corresponding microphone output signals, the reference signal would first be filtered with an estimate of this "error path". Such a pre-filter can be derived from a transfer function measurement made between the output of K and the microphone outputs when V_s is replaced with broadband noise, while the Accelophone™ is held up to a user's mouth without the user talking.

A preliminary simulation of this approach using transient signals such as those found in speech, with a microphone spacing of 2 cm and a filter order of 32, resulted in an overall reduction of approximately 11 dB in the pressure sum and an overall increase of about 8 dB in the pressure difference. While these results are indicative of the performance that may be achieved using this approach, the actual performance cannot be truly evaluated until we have implemented optimization algorithm with hardware "in the loop".

<u>Construction of a prototype for testing</u>

In order to test the AcceloPhone™ for actual performance, we built an early prototype of the system that has functionality sufficient to determine both its sensitivity to ambient interfering sound, and the amount of sound radiated to surrounding areas. The frequency range is limited to that required for understandable speech, from about 200 Hz to 2000 Hz. Electronic signal processing is done using a digital signal processor (DSP) with an A/D and D/A card for the early prototype. This prototype of AcceloPhone™ is to be used to optimize the signal processing method and confirm acoustical performance.

The hardware information and processing algorithms obtained previously allowed a prototype of the AcceloPhone™ to be constructed for testing. This prototype was designed to be used without a shroud and/or windscreen if possible, but there will likely be applications where a shroud is necessary and acceptable. If a shroud is needed, one as small as possible is desirable. The microphones need to be small, as close together, and as close to the loudspeaker as possible, consistent with the need for a measurable phase difference in microphone outputs. In order to deal with the inevitable phase mismatch between moderately priced microphones, it may be desirable to reverse their locations using a swiveling holder. This technique allows for phase calibration and is used for paired microphones in acoustical impedance tubes when the microphones are not closely matched in phase.

In this prototype, the AcceloPhone™ microphone signals is sampled using an A/D board in a dedicated DSP. The signal for the loudspeaker is continuously updated and generated in the computer, and fed to a power amplifier using a D/A channel on the same board. The processing and board control software will use appropriate software for the board of choice.

The early prototype constructed with its various options, is needed to be tested for sensitivity to ambient noise and for privacy. Both of these measurements can be made in a reverberation room. For noise sensitivity, one turns the adaptation off and measures the system signal output to noise and signal combined, in third octave bands from 200 to 2000 Hz. For these experiments the talker is replaced by a small loudspeaker emitting broad band repeated coherent signals, such as pseudo-random noise or chirp signals, band limited to match the spectrum of speech. The background noise is reproduced by an independent noise generator placed in the reverberant room. The signal to noise ratio is then determined and expressed by the articulation index (AI).

[redacted]

Functional Advantages of the Accelophone™

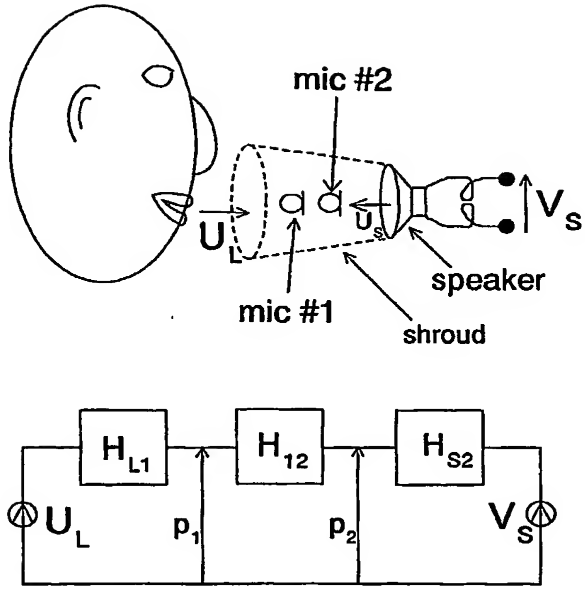
Immunity from ambient noise: The relative advantage of using air acceleration rather than pressure to improve directivity was shown in Figure 3. This directivity is for a "free field" sensor and does not take into account the effects of diffraction around the loudspeaker and talker, nor does it account for the effects of any shroud that may be employed. Those effects can be considered in modelling the system using the boundary element computations. When that calculation has been done, it will likely happen that there will be better microphone arrangements that reduce the sensitivity to ambient sound even more. Repositioning the microphones, increasing their number, or modifying the gain characteristics of the signal conditioning may be desirable.

A major decision point in the development will be whether or not a shroud of some type is needed for AcceloPhone™. We might think of the shroud as somehow "blocking" the ambient sound from arriving at the pair of microphones. But the shroud we have in mind and illustrated in Figure 4 is too small to do much blocking of ambient sound. The effect of the shroud is more likely to enhance the channeling of airflow from the talker "source" to the loudspeaker "sink" and therefore to provide larger air acceleration signal to be sensed.

Maintaining talker privacy and reducing bystander annoyance: As noted in this discussion, the source of sound for leakage away from the AcceloPhone™ is the total volume velocity due to both the talker and the loudspeaker. These sources are close in location, but not identical in location, to the microphone pair that senses the disturbance from ambient sound. Therefore there is not perfect reciprocity between the immunity from ambient sound and sound radiation away from the AcceloPhone™, especially at higher frequencies (between 2 and 10 kHz) where the wavelengths of sound approach the order of AcceloPhone™ dimensions. This means that optimization for immunity from ambient sound and optimization for privacy may not be totally consistent over the entire frequency range of interest.

Since the longer wavelength (lower frequency) sounds between 200 and 2000 Hz are of concern for noise immunity and the concern for privacy may concentrate on slightly higher frequencies (1500-5000 Hz), a choice of processing algorithm to deal with both may not be too difficult. Nevertheless, this aspect of system performance must be investigated as we proceed. A demonstration of this system using the prototype in Figure 2 showed a significant reduction in leakage sound and a simultaneous increase in signal from the microphone pair.

[redacted]



V_S is selected so that $p_1 \approx -p_2$

Fig. 1: System sketch

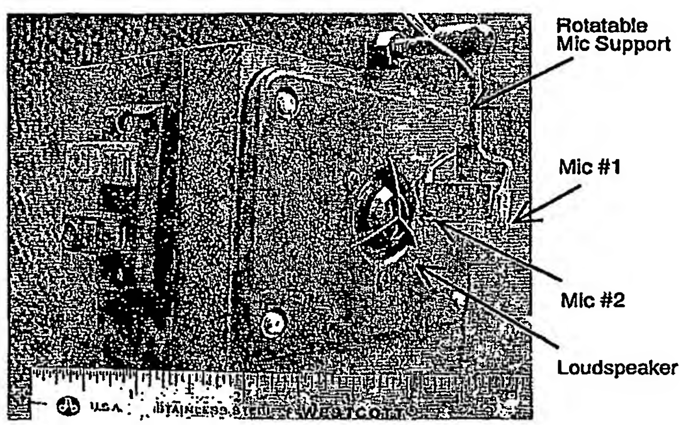


Fig. 2: Photograph of "Mock-up" of AcceloPhone™ set up for testing

[redacted]

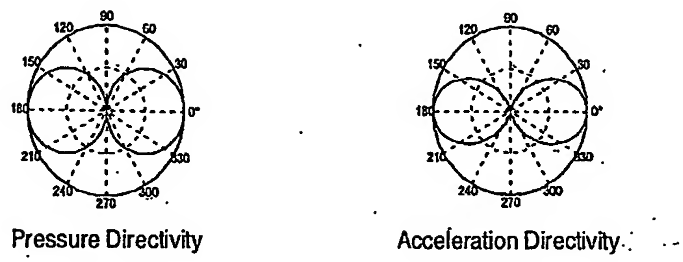


Fig. 3: Relative directivities for a sensor based on pressure and on acceleration

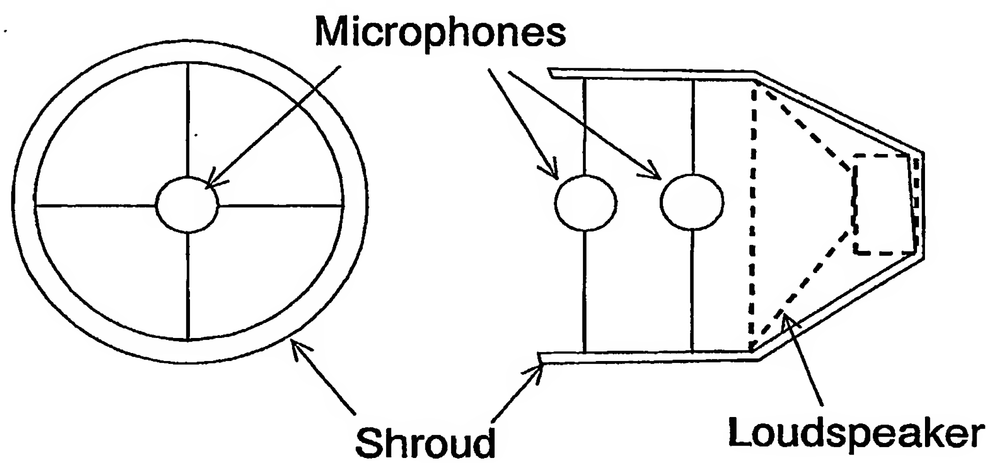


Fig. 4: Diagram of proposed transducer, showing microphones, speaker and housing layout

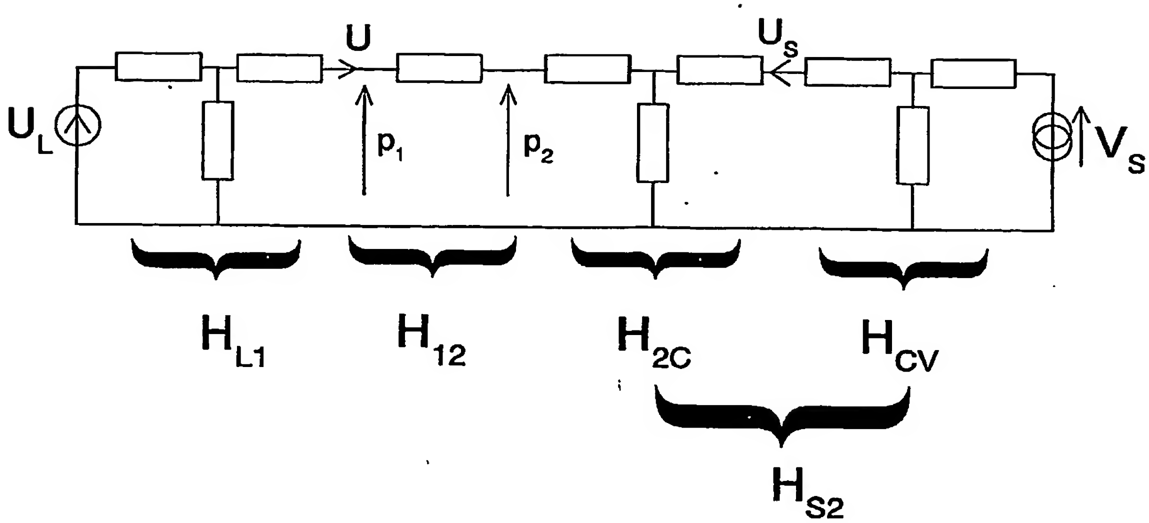


Fig. 5: Form of analytical model for proposed transducer

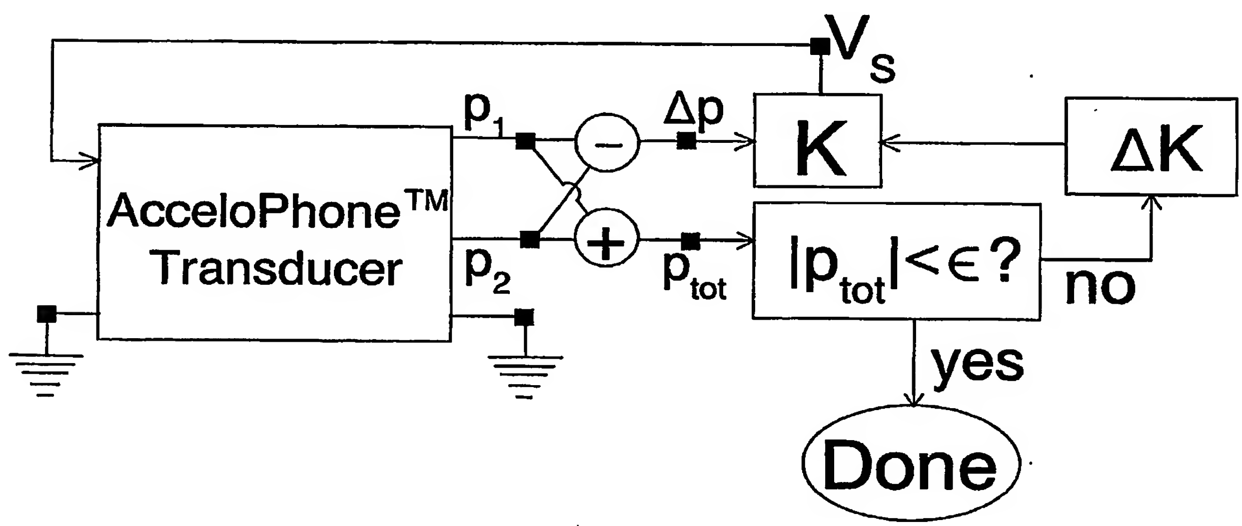


Fig. 6: Conceptual layout of optimization algorithm

[redacted]

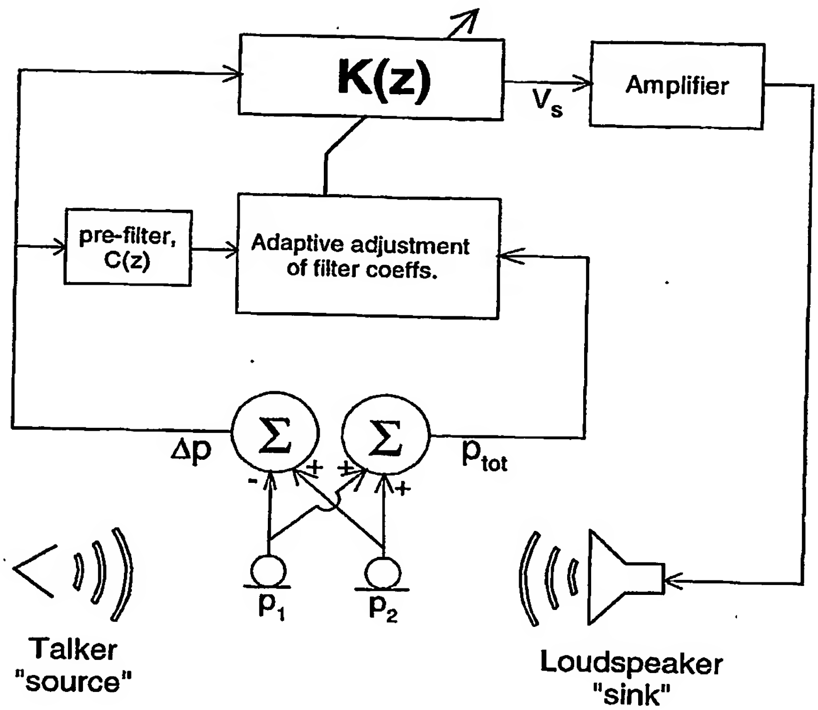


Fig. 7: Block diagram representing a general processing scheme for utilizing an adaptive controller to minimize p_{tot} while maximizing Δp

[redacted]

July 27, 2004

PCT DIVISION
34 CHEIMIN DES COLUMBETTES
GENEVA, 1211
GENEVA, 20, SWITZERLAND

This is to inform you that the document requested in your order dated 07-22-2004 is a true reproduction of the official office record copy of that document:

60430802 CERTIFIED PCT APP AS FILED-MAIL 1

☐ The enclosed office document is a reproduction from the official office record copy and has been recorded using high quality scanning or microfilm equipment. Copies of page/papers that were not scannable have not been included, nor have pages/papers received after the original filing date. Copies of these pages/papers may be ordered separately.

☑ The enclosed document is a reproduction of the best available source of the official office record copy of that document.

If you have any questions or need additional information, please contact our Customer Service Department.

Mailing Address:
Mail Stop Document Services,
Director of the U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Delivery Address:
U.S. Patent and Trademark Office
Office of Public Records, Customer Service
1213 Jefferson Davis Highway, Suite 300
Arlington VA 22202

Order your copies of USPTO documents on the Web. It's fast, easy, and secure! Go to (http://ebiz1.uspto.gov/oems25p/index.html) and order today!

Voice: (703) 308-9726 Fax: (703) 308-7048 E-Mail: mailto:DSD@USPTO.GOV

Ref:TL 1199942